

**МЕТОДИ ТА ЗАСОБИ ВИЯВЛЕННЯ НЕАКТУАЛЬНОЇ ТА НЕДОСТОВІРНОЇ  
ІНФОРМАЦІЇ НА WEB-РЕСУРСАХ****А.М. Мельник**Західноукраїнський національний університет,  
вул. Львівська, 11, м. Тернопіль, 46020, Україна, e-mail: melnyk.andriy@gmail.com

У статті розглянуто важливу науково-прикладну задачу виявлення неактуальної та недостовірної інформації на веб-ресурсах, яка є актуальним напрямком розвитку та реалізації методів інтелектуального аналізу даних. Проведено аналіз сучасних методів та засобів оцінки неактуальної та недостовірної інформації з точки зору оцінки інформаційних джерел та виділено основні проблемні напрямки, які виникають в процесі їх функціонування. Запропоновано систему показників фільтрування недостовірної та неактуальної інформації, яка отримана на основі кількох джерел. На основі даної системи реалізовано метод перевірки інформації з веб-ресурсів на актуальність та достовірність. В основі даного підходу покладено можливість використання наперед визначеного ресурсу, дані з якого носять тільки достовірний характер. Розроблено метод виявлення недостовірної та неактуальної інформації, враховуючи особливості її поширення через відповідні сторінки у соціальних мережах та використання багатозадачної класифікації інформації, яка отримується із різних джерел даних. Використання запропонованих інтелектуальних методів обробки даних разом із іншими методами інтелектуального аналізу, які використовуються для оцінки інформації отриманої з інтернету, дозволить значно підвищити ефективність процесу встановлення неактуальності та недостовірності інформації, а також дозволить побудувати оцінку конкретного веб-ресурсу на предмет публікації та розповсюдження такої інформації.

**Ключові слова:** актуальність та достовірність інформації, web-ресурс, інформаційне джерело, соціальні мережі.

**Вступ**

Виявлення неактуальної та недостовірної інформації є важливим елементом в процесі інформаційної обізнаності та прийняття. Велика кількість інформації, яка розповсюджується через Інтернет, вимагає її всебічної оцінки, а це в свою чергу породжує також питання оцінювання відповідного інформаційного джерела [1-3].

На відміну від друкованих матеріалів, таких як книги та статті, веб-сторінки не повинні відповідати стандартам якості, точності та статистичної достовірності. Ця відсутність контролю якості створює особливі проблеми для встановлення відповідного ступеня довіри [4], в межах якого необхідно оцінювати кожен веб-сторінку на предмет точності, достовірності, актуальності та об'єктивності [5].

В Інтернеті кожна адреса веб-сайту має домен як частину адреси, що дозволяє ідентифікувати власника веб-сайту. Домен або його приналежність до довіреної зони може бути швидким способом оцінити якість веб-ресурсу перед його відвідуванням. Як правило, .edu, .gov та .org частіше мають якіснішу інформацію, ніж домени .com [5-7].

У цій роботі реалізується процедура оцінки достовірності та актуальності інформації із конкретного веб-ресурсу. Застосування такої процедури також дозволяє провести загальну оцінку відповідного інформаційного джерела на предмет достовірності та актуальності.

## **Аналіз досліджень та публікацій**

Більшість публікацій з даної тематики зорієнтовані на використання методів інтелектуального аналізу [2,3,5,8], а саме штучних нейронних мереж, дерева рішень, використання символічних правил при побудові результируючих характеристик, правила, підходи на основі методів найближчого сусіда, методу опорних векторів, байєсових мереж, лінійної регресії, методів кореляційно-регресійного аналізу; методів кластерного аналізу, зокрема ієрархічні та неієрархічні методи кластерного аналізу [8]; методи пошуку асоціативних правил, метод Apriori; методи обмеженого або часткового перебору перебору, генетичні алгоритми в рамках еволюційного програмування, різноманітні методи візуалізації даних [2]. Більшість перерахованих методів були реалізовані в рамках штучного інтелекту [4-7]. В межах інтелектуального аналізу даних виникає кілька ключових задач, які вимагають розв'язку, виходячи із специфіки предметної області, в якій вони сформовані [7-9]. До таких задач треба віднести такі: класифікації, кластеризації, прогнозування, асоціація, візуалізації, аналіз виявлення відхилень, оцінювання, аналіз зв'язків, підведення підсумків [2,4,7-11].

Водночас, специфіка інформації, яка накопичується в інтернеті, або отримується з нього, вимагає реалізації додаткових підходів, які можна буде легко реалізувати та програмно інтерпретувати в рамках заданої предметної області [12-14]. Особливо актуальним є напрямок розробки методів аналізу інформації на предмет актуальності та достовірності [8].

### **Мета і задачі дослідження**

Отже, метою даної роботи є розробка методів та засобів виявлення неактуальної та недостовірної інформації, яка розміщується на офіційних веб ресурсах організацій та поширюється через відповідні соціальні мережі.

Для досягнення поставленої мети потрібно вирішити такі задачі:

- проаналізувати особливості процесу виявлення неактуальної та недостовірної інформації на веб-ресурсах;
- обґрунтувати використання метрики оцінки достовірності інформації для оцінювання інформації, яка отримується із веб-ресурсів;
- розробити процедуру перевірки інформації з веб-ресурсів на актуальність та достовірність;
- реалізувати метод перевірки інформації з веб-ресурсів на основі аналізу даних, отриманих із соціальних мереж;
- провести експериментальні дослідження та оцінити отримані результати на прикладі веб-ресурсу закладу вищої освіти.

### **Основна частина**

Перевірка веб-ресурсів на наявність неактуальної та недостовірної інформації вручну вимагає великих часових витрат. Така перевірка недостовірної інформації не масштабується за обсягом новоствореної інформації, особливо коли певна організація має кілька ресурсів, включаючи ресурси в різних соціальних мережах. Використання засобів, які спрямовані на автоматизацію такої перевірки, або часткове використання методів автоматичної перевірки в основному ґрунтуються на автоматизації процесів пошуку інформації, методів обробки природної мови, машинного навчання, графової теорії [5,6,11].

У загальному, представлену актуальну інформацію можна моделювати наступним відношенням [1,3,5]:

$$K_w = \langle S, P, O \rangle, \quad (1)$$

де  $K_w$  - інформація, яка моделюється наступною трійкою:  $S$  - предметна область,  $P$  - предикат,  $O$  - об'єкт.

Наприклад, розглянемо наступний інформаційний блок на веб-ресурсі класичного закладу освіти: «Західноукраїнський національний університет здійснює підготовку спеціалістів за спеціальністю інженерія програмного забезпечення». В даному блоці можна виділити наступні інформаційні блоки:

$$K_w = \langle \text{ЗУНУ}, \text{спеціальність}, \text{Інженерія програмного забезпечення} \rangle.$$

Більшість відомих засобів автоматичного видобування інформації ґрунтуються на представлені (1) та його відповідній інтерпретації.

Для пошуку неактуальної або недостовірної інформації та побудови відповідно системи її класифікації дані спочатку витягуються з відкритих джерел, які відносяться до однієї категорії або групи відповідних веб-ресурсів деякої установи чи організації. Такий процес часто відноситься до видобування даних або відповідних відношень між ними. Видобування даних можна класифікувати на видобуток даних з одного джерела або з відкритим кодом. Одне джерело видобування даних в основному ґрунтується на порівняно надійному веб-ресурсі (наприклад, офіційна веб-сторінка установи). Даний спосіб для отримання релевантної інформації є відносно ефективний, але часто призводить до встановлення неповної інформації, оскільки залежить від рівня наповнюваності або актуальності розміщеного контенту.

Пошук актуальної інформації на основі даних із кількох відкритих джерел є менш ефективним, але дозволить значно розширити формування повних та достовірних інформаційних блоків. Для пошуку неактуальної та недостовірної інформації необхідно також сформувані наступні множини (2), (3), (4):

$$A_t \in K_w, \quad (2)$$

де  $A_t$  - представляє фрагмент даних, який вдалося верифікувати.

$$\text{Base}A_t = \{A_t\}, \quad (3)$$

де  $\text{Base}A_t$  - представляє множину верифікованих даних.

$$\text{BaseRelation} = \{\text{Base}A_t, \text{Relation}A_t\}, \quad (4)$$

де  $\text{BaseRelation}$  - представляє множину зв'язків між сутностями, які описані за допомогою відповідних відношень  $\text{Relation}A_t$ .

Для того, щоб сформувані базу із достовірною інформацією, яка отримана на основі кількох джерел, їх потрібно додатково відфільтрувати, враховуючи наступні особливості:

- Надлишковість представлених даних в одному контексті (наприклад,  $K_w = \langle \text{ЗУНУ}, \text{спеціальність}, \text{Інженерія програмного забезпечення} \rangle$ ) та

$$K_{w_1} = \left\langle \begin{array}{l} \text{Західноукраїнський національний університет, спеціальність,} \\ \text{Інженерія програмного забезпечення} \end{array} \right\rangle$$

$Kw_i$  буде надлишковим представленням, оскільки «ЗУНУ» та «Західноукраїнський національний університет» відповідають одній і тій же сутності;

- Недійсність – показник, який залежить від конкретного інтервалу часу, наприклад,  $Kw = \langle THEU, \text{університет}, \text{Україна} \rangle$  неактуальна інформація і повинна бути оновлена. Один із варіантів подолання даної проблеми – представлення даних із певним набором часових характеристик або розширення отриманого набору за допомогою встановлення додаткових тверджень;

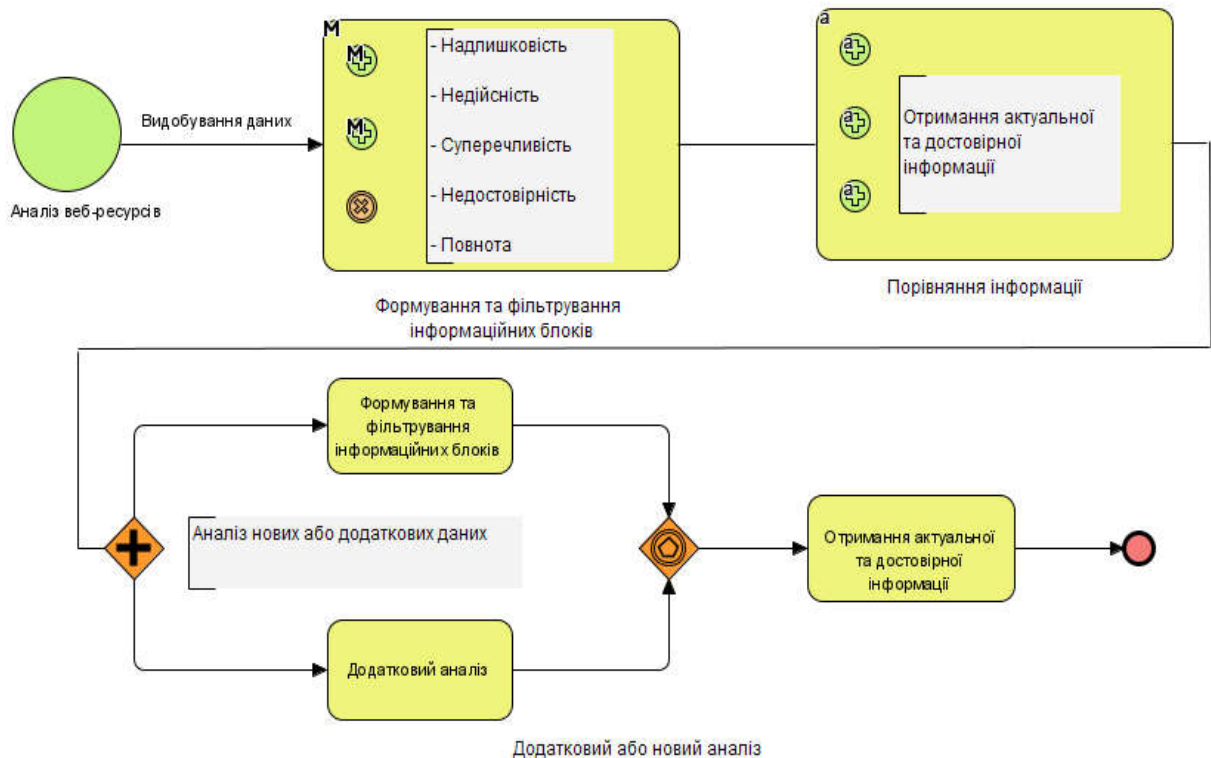
- Суперечливість представлених даних (наприклад  $Kw = \langle ЗУНУ, \text{адреса}, \text{Тернопіль} \rangle$  та  $Kw = \langle ЗУНУ, \text{адреса}, \text{Вінниця} \rangle$  є суперечливими даними, які вимагають додаткового аналізу);

- Недостовірність веб-ресурсу (складність ідентифікації ресурсу до довіреної зони в рамках заданої предметної області). Достовірність веб-ресурсу можна встановити наприклад через приналежність його доменної адреси до довіреної зони;

- Повнота – наявність інформації, яка отримана із веб-ресурсів не завжди дозволяє встановити її достовірність та актуальність.

**Процедура перевірки інформації з веб-ресурсів на актуальність та достовірність.** Щоб оцінити достовірність інформації, яка розміщується на офіційних веб-ресурсах установ та організацій, необхідно порівняти дані, які отримані та представлені за допомогою відношення (1) з інформацією, яка отримана з достовірних джерел.

Як правило, процедура перевірки даних для трійки  $Kw = \langle S, P, O \rangle$  полягає в оцінці можливості того, що межі, позначені предикатом  $P$ , сформовані із врахуванням приналежності  $S$ , до вузла, що представляє  $O$  в множині  $BaseAt$ . Зокрема, цей процес можна описати за допомогою наступних кроків, які формалізовано представлені на рисунку 1.



**Рис. 1.** Схема аналізу інформації з веб-ресурсів із використанням показників оцінки достовірності інформації.

1. Розташування сутності. Тема  $S$ , так само як і об'єкт  $O$  спочатку повинні відповідати елементу із множини достовірних даних в  $BaseAt$ , який представляє ту саму сутність.

2. Перевірка відношень. Інформація, яка описується за допомогою (1) вважається актуальною та достовірною, якщо межа позначене предикатом від елемента, що представляє  $S$ , до елемента, що представляє  $O$ , існує в множині опису зв'язків  $BaseRelation$ . В іншому випадку інформація вважається хибною, або такою що потребує додаткових уточнень.

3. Встановлення достовірності інформації. Коли  $Kw = \langle S, P, O \rangle$  не входить у множину  $BaseRelation$ , ймовірність для меж сутності, яка позначена предикатом в межах певної предметної області можна визначити, використовуючи показник семантичної близькості [5], який дозволяє провести попарну оцінку елементів із множини  $\langle S, P, O \rangle$ , які знаходяться між собою в семантичних відношеннях (синоніми, гіперо-гіпонімічні відношення, асоціативність), і нульові значення для всіх інших пар.

Розглянемо специфіку описаного процесу перевірки. Припустимо, що верифікований на офіційному веб-ресурсі інформаційний блок можна представити як формалізований набір тверджень  $Kw = \langle s_i, p_i, o_i \rangle, i = 1, 2, \dots, n$ . Множина  $W_{BaseAt}$  сформована на основі верифікованих наборів даних  $st_j, pt_j, ot_j, j = 1, 2, \dots, m$ .

Перевірка інформаційних блоків, які розміщені на суміжних до досліджуваного веб-ресурсу буде зводитися до побудови функцій  $Fa$  із знайденими показниками автентичності  $Af_i \in [0; 1]$  для кожного блоку  $\langle s_i, p_i, o_i \rangle$ , попарно їх порівнюючи із  $\langle st_j, pt_j, ot_j \rangle$  із врахування опису відношень із  $BaseRelation$ :

$$\begin{aligned} Af_i &= 1, \text{ якщо блок верифікований як достовірний;} \\ Af_i &= 0, \text{ якщо блок верифікований як недостовірний.} \end{aligned}$$

Загальний показник автентичності ресурсу  $Af$  буде визначатися як агрегована величина усіх  $Af_i$ :

$$Fa : (s_i, p_i, o_i) \xrightarrow{W_{BaseAt}} Af_i \quad (5)$$

$$Af = O(Af_i), \text{ де } i = 1, \dots, n, \quad (6)$$

де  $O$  – вибрана функція агрегування (наприклад, середньозважена).

Інформація, яку необхідно перевірити, є достовірною при умові  $Af = 1$ , а при  $Af = 0$  вказана інформація повністю недостовірною:

$$Fa((s_i, p_i, o_i), W_{BaseAt}) = P(\text{межа } p_i, \text{ що зв'язує } s_i \text{ та } o_i \text{ в } W_{BaseAt}), \quad (7)$$

де  $P$  – ймовірність відповідності  $s_i, o_i$  до достовірних даних  $st_j, ot_j$  в  $W_{BaseAt}$ .

$$s_i = \underset{st_j}{\operatorname{argmin}} [U(s_i, st_j)] < \theta, \quad (8)$$

$$o_i = \underset{o_{t_j}}{\operatorname{arg\,min}} [U(o_i, o_{t_j})] < \theta, \quad (9)$$

де  $U$  – відстань між двома елементами відповідних сутностей, яку можна розрахувати за допомогою формули Джакардової відстані[5,12,13]. Величина  $U$  буде відображати ідентифікацію двох сутностей між собою, наприклад, якщо  $U(o_i, o_{i+1}) = 0$ , то  $o_i$  та  $o_{i+1}$  ті самі сутності.

У результаті отримаємо множину видобутих даних та множину верифікованих достовірних даних. Однією із основних проблем при побудові множини достовірних даних буде джерело (джерела), з якого ці дані можна отримати. В якості відповідного джерела може використовуватися наперед визначений ресурс, дані з якого несуть тільки достовірний характер.

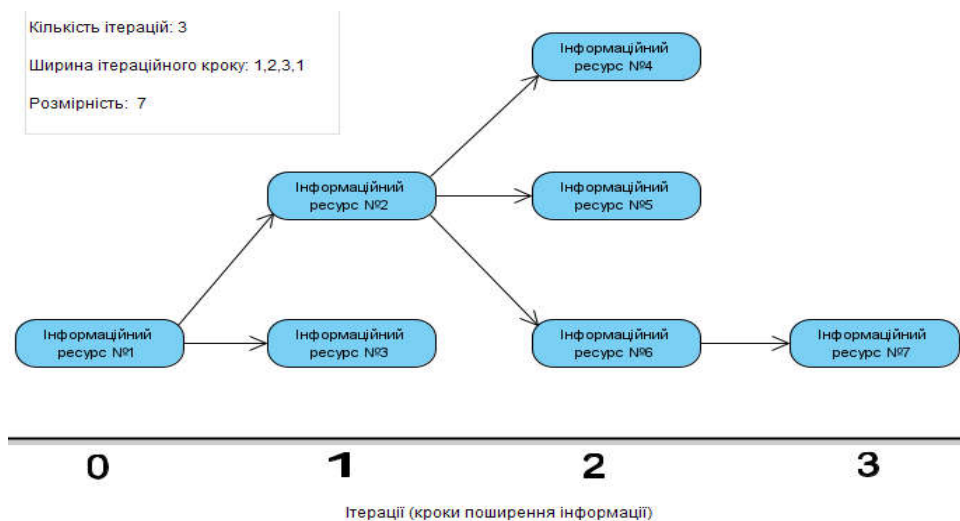
**Метод перевірки інформації з веб-ресурсів на основі аналізу даних, отриманих із соціальних мереж.**

Виявлення недостовірної інформації, враховуючи особливості її поширення через відповідні сторінки у соціальних мережах, можна реалізовувати, використовуючи методи аналізу структури інформації в процесі їх поширення або дублювання. Виявлення недостовірної інформації на основі її розповсюдження можна сформулювати як задачу багатозадачної класифікації інформації, яка отримується із різних джерел даних. Вхідними даними даного методу поширення може бути впорядкована множина, елементами якої є інформаційні блоки та відповідні показники їх ідентифікації (час, джерело, профіль).

Таку впорядковану множину можна представити через деревоподібну структуру, яка, враховуючи відповідні показники, фіксує ступінь поширення певної інформації через соціальні мережі. Кореневий елемент такої структури визначає джерело, з якого інформація була взята в початковому вигляді, а інші вузли представляють користувачів, які поширювали інформацію, враховуючи їх взаємозв'язок із батьківським джерелом (наприклад, офіційна сторінка організації в соціальній мережі та сторінки користувачів, які є учасниками відповідної групи або просто «друзями»).

Оцінку достовірності інформації, яка поширюється через відповідні профілі у соціальних мережах, можна здійснювати через врахування показника потужності множини або через їх оцінку через часові характеристики.

На рисунках 2 та 3 наведено ілюстрацію оцінки достовірності інформації, яка отримується через соціальні мережі із використанням зазначених вище методів оцінки.



**Рис. 2.** Оцінка достовірності інформації на основі ітераційної оцінки множин

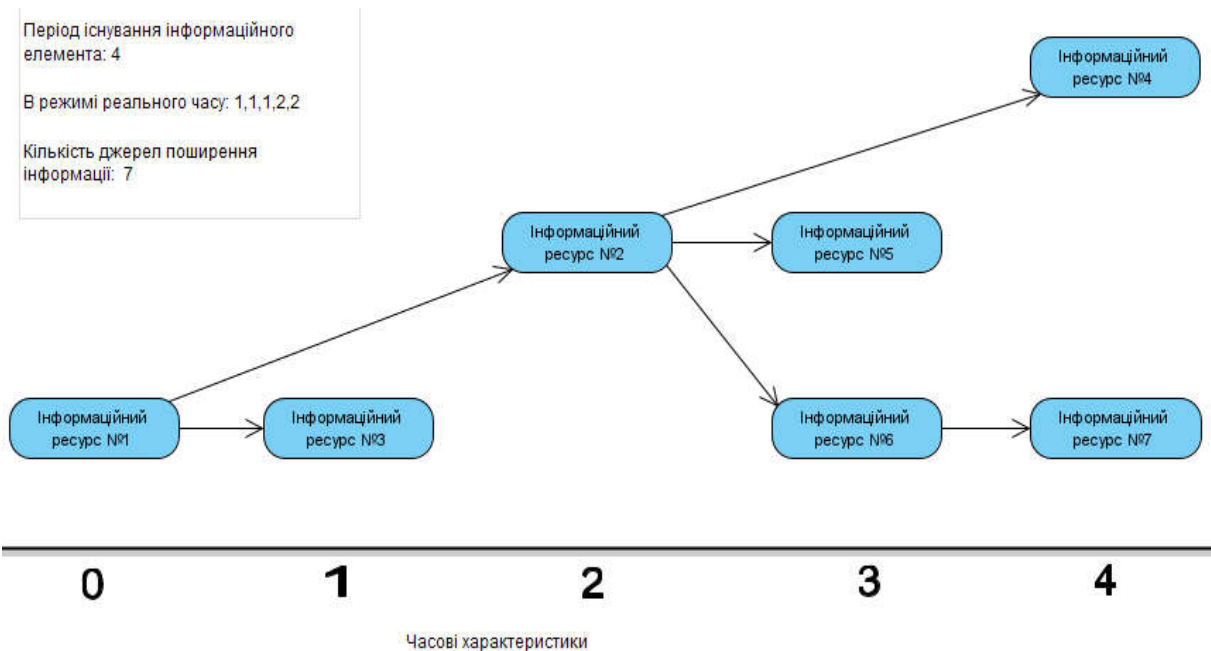
Основними критеріями, які використовуються при використанні ітераційної оцінки множин, є:

- кількість ітерацій - максимальна кількість ітерацій, на яких відбувалося поширення інформаційних блоків;
- ширина ітераційного кроку – кількість джерел інформації в межах ітераційного кроку;
- розмірність – загальна кількість джерел інформації, через які відбувається поширення;

На рисунку 3 відображено ці показники, виходячи із побудованої деревовидної структури.

Основними критеріями, які використовуються в процесі використання часової оцінки, є:

- період існування інформаційного елемента—час, протягом якого інформація поширювалася через відповідні інформаційні джерела;
- режим реального часу – кількість джерел інформації, які поширювали інформацію в певний час;
- кількість джерел поширення інформації – загальна кількість джерел інформації, через які відбувається поширення;



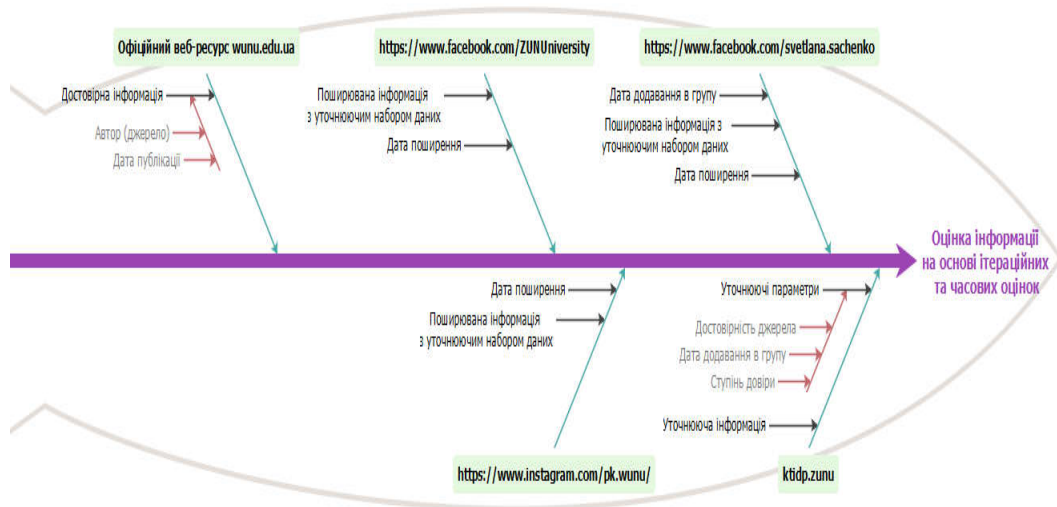
**Рис. 3.** Оцінка достовірності інформації на основі часової оцінки множин

Складність застосування такого методу визначення достовірної інформації полягає у розширенні кількості вузлів на кожному етапі ітерації, наприклад, через велику чисельність активних користувачів, які належать до кореневого вузла. Крім того, часто доводиться фільтрувати додаткові показники, які можуть нарощуватися на кожній ітерації. В загальному випадку достовірність інформації буде визначатися, виходячи із її близькості до кореневого вузла поширення із врахуванням актуальних часових характеристик.

### Експериментальні дослідження

На основі запропонованого підходу було проведено ряд експериментальних досліджень. Суть даних досліджень зводилася до оцінки інформації веб-ресурсу класичного університету Тернополя (<https://www.wunu.edu.ua/>) та відповідних джерел

інформації, які до нього повністю, частково або не прив'язані. На рисунку 4 представлено схему, яка відображає результати оцінки відповідних джерел інформації із врахуванням параметрів на кожному ітераційному кроці.



**Рис. 4.** Оцінка достовірності інформації на основі ітераційних та часових характеристик на прикладі даних, отриманих з ресурсу <https://www.wunu.edu.ua/>

В результаті експертного оцінювання з елементами використання автоматизованих інструментальних засобів аналізу контенту було підтверджено достовірність інформації, яка відображалася на кожному кроці оцінювання. Відповідні результати наведено у таблиці 1.

Оцінка достовірності здійснювалася на основі агрегованої оцінки кожного інформаційного джерела. В результаті встановлено, що на трьох інформаційних ресурсах, інформація, яка поширюється, може вважатися достовірною, а на двох - частково достовірною.

**Таблиця 1.**

Результати оцінювання достовірності інформації на основі поширення її через соціальні мережі в рамках структури відношень домена <https://www.wunu.edu.ua/>

Параметри оцінювання	<a href="https://www.wunu.edu.ua/">https://www.wunu.edu.ua/</a> (достовірне джерело)	<a href="https://www.facebook.com/ZUNUniversity">https://www.facebook.com/ZUNUniversity</a>	<a href="https://www.facebook.com/svetlana.sachenko">https://www.facebook.com/svetlana.sachenko</a>	<a href="https://www.instagram.com/pk.wunu/">https://www.instagram.com/pk.wunu/</a>	ktidp.zunu
Автор	Прес-центр	Прес-центр	Саченко	Приймальна комісія	Кафедра історії держави
Дата публікації (поширення)	20.01.2021	20.01.2021	21.01.2021	21.01.2021	20.01.2021
Достовірність джерела	Достовірне	Достовірне	Частково достовірне	Достовірне	Достовірне
Дата реєстрації джерела в достовірному домені	-	20.09.2020	31.10.2020	21.09.2020	30.09.2020
Ступінь довіри	Високий	Високий	Середній	Високий	Середній
Оцінка достовірності	Достовірна	Достовірна	Частково достовірна	Достовірна	Частково достовірна



## Висновки

У роботі розглянуто особливості процесу оцінки інформації на предмет актуальності та достовірності. Відсутність належних підходів, які прості у використанні та реалізації, спонукали до розробки відповідної системи показників та методів оцінки недостовірної та неактуальної інформації. Використання запропонованих в рамках даної роботи методів дозволило підвищити ступінь розпізнавання джерела інформації на предмет розміщення неактуальної та недостовірної інформації (Таблиця 1.)

Створені методи програмно реалізовані як надбудова до існуючих рішень в рамках окремих програмних модулів і можуть ефективно використовуватися для оцінки достовірності інформації. Ефективність запропонованих методів підтверджена експериментально на прикладі оцінки веб-ресурсів закладу вищої освіти - Західноукраїнського національного університету, а також пов'язаних із ним інформаційних ресурсів.

## Список літератури

1. Vo N., Lee K. The Rise of Guardians: Fact-checking URL Recommendation to Combat Fake News. *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 2018, P. 275–284. URL: <https://doi.org/10.1145/3209978.3210037>
2. Колодчак О.М. Інтелектуальний аналіз даних. *Вісник Національного університету "Львівська політехніка". Комп'ютерні системи та мережі*. 2013. № 773. С. 49-58.
3. Ye J., Skiena S. Media Rank: Computational Ranking of Online News Sources. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining July*. 2019. P. 2469–2477 URL: <https://doi.org/10.1145/3292500.3330709>.
4. Yu B., Zhang Z., Tingwen L., Wang B., Li S., Li Q. Beyond Word Attention: Using Segment Attention in Neural Relation Extraction. *IJCAI*. 2019. P. 5401-5407.
5. Chao W. Relation Extraction. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization*, 5401–5407. URL: <https://doi.org/10.24963/ijcai.2019/750>.
6. Xinyi Z., Reza Z. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Comput. Surv.* V.53, N5, Article 109. 2020, 40 p. DOI: <https://doi.org/10.1145/3395046>.
7. Ray O., Qian J., Wang W.Y. A Survey on Natural Language Processing for Fake News Detection. *ArXiv abs/1811.00770* (2020): n. pag.
8. Dyvak M., Papa O., Melnyk A., Pukas A., Porplytsya N., Rot A. Interval Model of the Efficiency of the Functioning of Information Web Resources for Services on Ecological Expertise. *Mathematics*, 2020. V. 8 No. 12, 2116. URL: <https://doi.org/10.3390/math8122116>.
9. Nørregaard J., Horne B.D., Adalı S. A Large Multi-Labelled News Dataset for the Study of Misinformation in News Articles. *Proceedings of the International AAAI Conference on Web and Social Media*, 2018. V. 13., P. 630–638.
10. Trivedi R., Sisman B., Dong, X., Faloutsos C., Ma J., Zha H., Link N. Bed: Multi-Graph Representation Learning with Entity Linkage. 2018. P. 252-262. URL: [doi10.18653/v1/P18-1024](https://doi.org/10.18653/v1/P18-1024).

11. Vilone G., Longo L. Explainable Artificial Intelligence: a Systematic Review. 2020.
12. Kovbasistyi A., Melnyk A., Dyvak Brych V., Spivak I. Method for detection of non-relevant and wrong information M. based on content analysis of web resources. *XIII International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH)*. Lviv, 2017, P. 154-156, URL: doi: 10.1109/MEMSTECH.2017.7937555.
13. Dyvak M.P., Kovbasistyi A.V., Melnyk A.M., Turchyn L.Y., Martsenyuk Y.O. System for web resources content structuring and recognizing with the machine learning elements. *Radio Electronics, Computer Science, Control*, 2018. N 3. <https://doi.org/10.15588/1607-3274-2018-3-14>.
14. Дивак М.П., Мельник А.М., Ковбасистий А.В., Папа О.А. Підхід до математичного моделювання ефективності web-ресурсів. *Оптико-електронні інформаційно-енергетичні технології*. 2020. Т. 38, N. 2. P. 29–37. URL: <https://doi.org/10.31649/1681-7893-2019-38-2-29-37>.

#### METHODS AND TOOLS OF DETECTING CURRENT AND RELIABLE INFORMATION ON WEB-RESOURCES

A.M. Melnyk

West Ukrainian National University,  
11, Lvivska Str., Ternopil, 46020, Ukraine; melnyk.andriy@gmail.com

The article considers an important scientific and applied task of identifying irrelevant and unreliable information on web resources, which is an important area of development and implementation of methods of data mining. The analysis of modern methods and means of estimation of irrelevant and unreliable information from the point of view of estimation of information sources is carried out and the basic problem directions which arise in the course of their functioning are allocated.

A system of indicators for filtering unreliable and irrelevant information, which is obtained on the basis of several sources, is proposed. Based on this system, a method of checking information from web resources for relevance and reliability has been implemented. This approach is based on the possibility of using a predefined resource, the data from which are only reliable.

A method of detecting inaccurate and irrelevant information has been developed, taking into account the peculiarities of its distribution through relevant pages in social networks and the use of multitasking classification of information obtained from various data sources.

The use of the proposed intelligent data processing methods together with other methods of intellectual analysis used to evaluate information obtained from the Internet, will significantly increase the efficiency of the process of establishing irrelevance and inaccuracy of information, and will build an assessment of a particular web resource for publishing and disseminating such information.

**Keywords:** relevance and reliability of information, web-resource, information source, social networks.